J. N. Am. Benthol. Soc., 2008, 27(4):932–948

DOI: 10.1899/08-041.1
Published online: 28 October 2008

Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US

Alan T. Herlihy[1]

Department of Fisheries and Wildlife, Oregon State University, Corvallis, Oregon 97331 USA

Jean C. Sifneos[2]

Department of Geosciences, Oregon State University, Corvallis, Oregon 97331 USA

Abstract. We analyzed nutrient data from a probability survey of 1392 wadeable streams across the 48 conterminous states of the US and from intensified survey data in 921 streams in the Pacific Northwest (PNW) to examine different methods of setting nutrient criteria and to develop a nutrient stream typology. We calculated potential nutrient criteria for total P (TP) and total N (TN) by 3 methods (ecoregion population 25th percentile of population, least-disturbed reference-site 75th percentile, and disturbance modeling) and compared them with existing draft US Environmental Protection Agency (EPA) criteria within 14 national nutrient ecoregions. All criteria derived from the methods were highly correlated; however, absolute values within ecoregions differed greatly among approaches. Population 25th percentiles of TP were almost always lower from statistically designed survey data than from found data. TN percentiles were more similar than were TP profiles, but they still tended to be lower from survey data than from found data. TP and TN population 25th percentiles were lower (often by a factor of 2–6) than reference-site 75th percentiles in all ecoregions. This result indicates that population 25th percentiles cannot be used as surrogates for reference-site 75th percentiles. Thirty-nine percent of the assessed national stream length exceeded TP criteria and 47% exceeded TN criteria when compared to nutrient criteria based on EPA Wadeable Stream Assessment reference-site 75th percentiles. In the PNW data set, all disturbance regression model estimates of background nutrient concentrations were lower than reference-site 75th percentiles. Regression tree analysis based on PNW reference sites used runoff, elevation, acid neutralizing capacity, forest composition, substrate size, and Omernik level III ecoregion as environmental class predictors to explain 46 to 48% of the total deviance in nutrient concentration. Reference-site nutrient concentrations varied widely among Omernik level III ecoregions in nutrient ecoregion II. Our analysis and the literature strongly suggest that 14 national nutrient ecoregions are too coarse to account for natural variation in stream nutrient concentrations. Setting appropriate national nutrient criteria will require finer-scale typology or classification of sites that better controls for natural variation.

Key words: nutrients, nutrient criteria, streams, typology, classification, reference condition, United States.

According to a 1996 report made by the US Environmental Protection Agency (EPA) to Congress, a large percentage of water bodies in the US have impaired water quality from excess nutrients (USEPA 1996). Excess nutrients alter aquatic trophic state and cause detrimental effects, such as noxious algal blooms, fish kills, and reduced water clarity. The relationship between nutrients and trophic state is direct in lakes because lakes are autotrophic, but this relationship is less direct in streams, which often are heterotrophic (Dodds 2007). Thus, understanding eutrophication and establishing nutrient criteria are important issues in streams.

The EPA has devised a strategy to develop regional nutrient criteria (USEPA 1998). Implicit in the strategy is recognition that excess nutrients are a major cause of water-quality impairment in the USA. The strategy also acknowledges that a single national nutrient criterion for all types of water bodies is inappropriate because of diverse geology, climate, and geomorphology. A series of technical guidance manuals released

[1] E-mail addresses: alan.herlihy@oregonstate.edu
[2] sifneos@science.oregonstate.edu

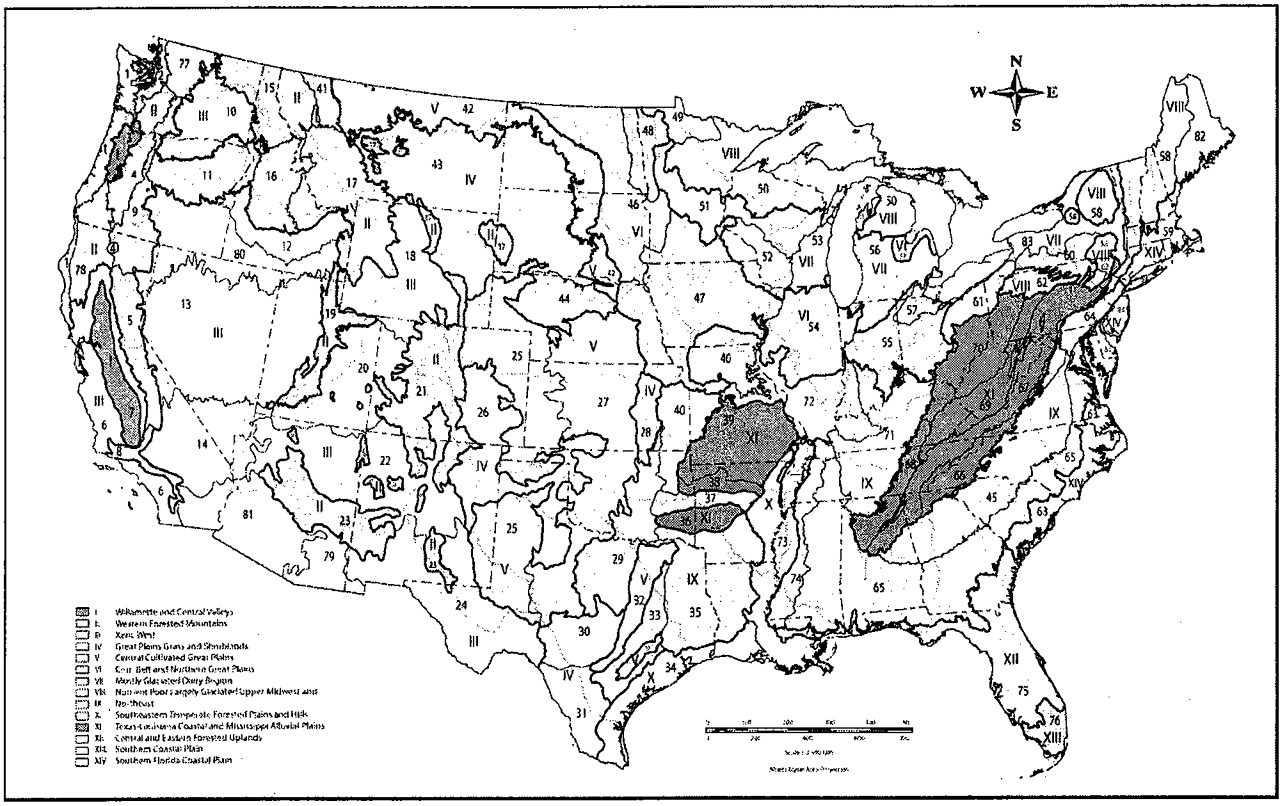

FIG. 1. National nutrient ecoregions aggregated from 84 Omernik level III ecoregions on the basis of geology, land use, ecosystem type, and nutrient conditions. Roman numerals refer to nutrient ecoregions, numbers indicate Omernik level III ecoregions within nutrient ecoregions.

by the EPA (USEPA 2000a) presented different approaches for developing nutrient criteria for different water bodies. For streams, approaches included use of reference reaches, downstream uses, best professional judgment, historical data, and predictive modeling.

The reference-reach approach is predicated on the idea that nutrient concentrations comparable with those in undisturbed streams would be protective of the designated uses of a stream or river. Thus, appropriate nutrient criteria can be estimated from conditions observed at relatively undisturbed reference sites. The objective of our paper is to examine potential reference-reach-based nutrient criteria for wadeable streams in the 48 conterminous states.

EPA conducted an analysis of historical found data for each of 14 national nutrient ecoregions (derived by aggregating 84 Omernik level III ecoregions [Omernik 1987] as described in Rohm et al. 2002; Fig. 1) and drafted preliminary nutrient criteria based on the 25th percentile of this found data (e.g., USEPA 2000b). However, analyzing data at national or even regional scales is difficult. Often data from multiple surveys done by many different groups must be combined, and doing so can be problematic because of differing protocols, study areas, database formats, survey objectives, data access, and site representativeness.

Between 2000 and 2004, EPA conducted the Wadeable Streams Assessment (WSA) using a probability sample of the wadeable streams in the 48 conterminous states. Use of WSA data for estimating nutrient criteria alleviates most of the problems with combining data because the data are recent (post-2000), collected with the same field protocol, and stored in similar data formats. Furthermore, the sample sites were picked in a systematic randomized manner so they are representative of the study area. Thus, the WSA data provide a unique opportunity to study stream classification and nutrient criteria at the national scale.

We used WSA data to quantify percentiles of nutrient distributions in all sites and in least-disturbed (reference) sites to estimate nutrient criteria for each nutrient ecoregion. Reference condition for streams is difficult to define because few data exist to tell us what

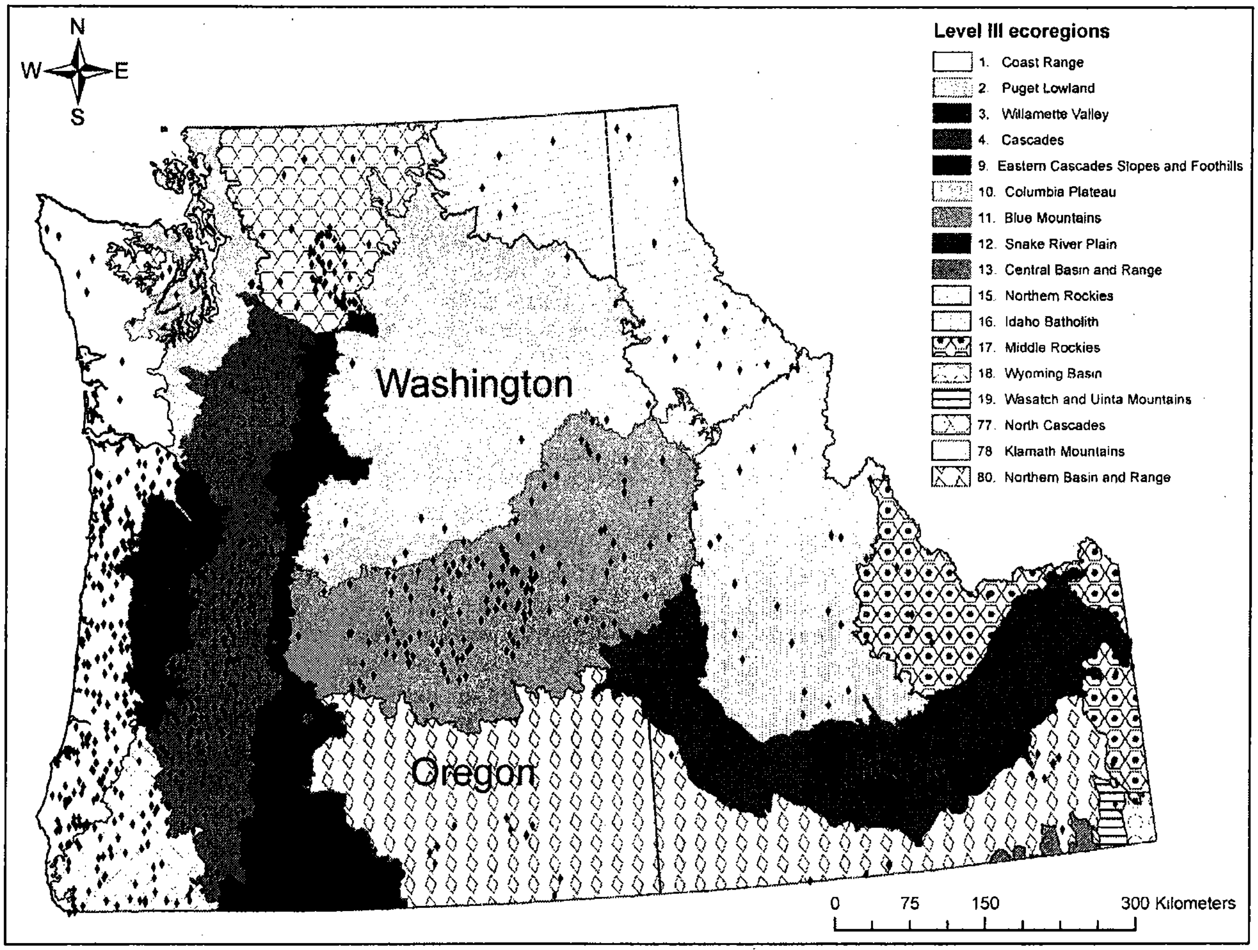

FIG. 2. Omernik level III ecoregions in the Pacific Northwest (PNW) and locations of PNW sampling sites.

undisturbed sites really looked like. In an ideal world, nutrient criteria would be based on observed nutrient concentrations in undisturbed systems. However, undisturbed systems do not exist anymore, and we must somehow approximate them either by modeling past conditions or by estimating undisturbed conditions from data collected from least-disturbed systems.

We also examined potential nutrient criteria at a finer scale using a more extensive data set collected in the Pacific Northwest (PNW; consisting of portions of nutrient ecoregions I, II, and III; Figs 1, 2), in which WSA data were augmented with other probability survey data from the region. We used the finer-scale PNW data set to examine the effect of the choice of screening criteria used to define least-disturbed reference conditions on reference-site nutrient concentration percentiles. We tested reference-site definitions based on water quality, aerial orthophotographs, biological metrics, and landscapes. We then used the PNW reference data set to examine variation within nutrient ecoregions. We also analyzed PNW reference sites to develop a regional stream classification scheme or typology to define groups of streams that would be expected to be similar with regard to reference-site nutrient concentrations and nutrient criteria.

Methods

Survey data sets

National.—WSA was a product of 2 surveys. In the 1st survey, flowing waters (streams and rivers) in the 12 western states were sampled during summer 2000 to 2004 as part of the EPA Environmental Monitoring and Assessment Program (EMAP) Western Pilot Study (EMAP-West) (Stoddard et al. 2005). The 841 EMAP probability sites that were wadeable (could be sampled safely by field crews wading the stream) were used in the WSA. In the 2nd survey, another 551 wadeable-stream probability sites were sampled in the

36 eastern and midwestern states in summer 2004. In both surveys, probability sites were selected using the randomized EMAP sampling design from the digital stream network depicted on 1:100,000 scale US Geological Survey (USGS) topographic maps (Herlihy et al. 2000, Stevens and Olsen 2004, Olsen and Peck 2008) to ensure that the samples were representative of the surveyed regions and that statistically valid population percentiles could be estimated. In addition, 333 hand-picked sites in the west and 143 hand-picked sites in the east were sampled in an attempt to augment the number of least-disturbed sites. We did not use hand-picked sites to make national population estimates, but we did use both hand-picked and probability sites as potential reference sites if they passed the reference screening criteria described in detail later.

PNW.—A number of regional EMAP surveys have been conducted in the PNW since 1994, and the Oregon Department of Environmental Quality (DEQ) has been sampling other streams in Oregon using EMAP methods. Sample sites in nutrient ecoregion I were sparse in these surveys, so we also obtained data from the EPA Agriculture–Riparian project (Moser et al. 1997) and a prepilot survey of streams conducted by Oregon State University (Herlihy et al. 1997). In all of these surveys, ~½ of the sample sites were selected with a systematic randomized sampling design (probability sites). We used only probability sites to estimate population percentiles. We used an existing Oregon DEQ reference-site database to increase our sample of potential reference sites. All together, we analyzed 921 wadeable stream sample sites (393 probability, 528 hand-picked) spread throughout the PNW (Fig. 2).

Data collection

All data were collected using field protocols developed for EMAP surveys (Peck et al. 2006). A single grab sample for water chemistry was taken at each site, and all field data were collected during a summer index period (June–September). Water samples were analyzed for nutrients including total P (TP), total N (TN), NO_3, and NH_4. TP was analyzed in the laboratory with persulfate digestion and colorimetry (USEPA 1987). TN was analyzed by persulfate digestion in all EMAP and WSA data, but it was calculated by summing total Kjeldahl N and NO_3 in Oregon DEQ data. Water samples also were analyzed for major anions and cations, conductivity, and turbidity.

Macroinvertebrate assemblages were sampled by compositing 8 to 11 kick net (500-µm mesh) samples. Macroinvertebrate bioassessment metrics related to stream ecological condition were calculated for all sites.

Physical-habitat data were collected from 11 equal-interval transects in each study reach. Physical-habitat measurements included substrate size, sinuosity, canopy cover, slope, riparian disturbance, and channel size (Kaufmann et al. 1999). Substrate size was quantified using a modified Wolman pebble count (Kaufmann et al. 1999).

Human disturbance in the riparian zone was assessed on both banks at each transect. Disturbance types included buildings, roads, pipes, landfill/trash, parks/lawn, pavement, revetment/walls, row crops, pasture/range, logging, and mining. Riparian disturbance at each transect was categorized as being on the bank (weight = 1.5), within a 10×10-m plot next to the transect (weight = 1.0), >10 m from the transect (weight = 0.667), or absent (weight = 0). A riparian disturbance score was calculated for each type of disturbance by taking a weighted average of the 22 observations for each type of disturbance (Kaufmann et al. 1999). For example, a score of 1.5, indicating buildings on both banks at each of the 11 transects, would be the highest possible value of the building riparian disturbance score (RDbuilding). The EMAP riparian disturbance index was calculated by summing the individual disturbance-type scores. Summary metrics for total agricultural riparian disturbances and nonagricultural riparian disturbances were calculated by summing appropriate individual disturbance scores. In addition, all watersheds were digitized and available geographical information system (GIS) data layers were used to calculate metrics for watershed land use/land cover, precipitation, runoff, air temperature, and elevation. In the PNW analyses for nutrient ecoregion I, we used a higher-resolution set of land use/land cover data derived from aerial photographs that was available from activities associated with the Willamette Futures project (Van Sickle et al. 2004) instead of GIS layers.

Calculations for candidate nutrient criteria

A variety of reference-reach methods for setting nutrient criteria have been proposed in the EPA technical guidance documents for streams (USEPA 2000a). We considered percentiles of the population of all streams, percentiles of reference streams, and models of natural background concentrations.

Reference-site 75th percentiles.—One approach calls for using the 75th percentile of values at reference sites as nutrient criteria. The EPA chose the reference-site 75th percentile because it is associated with minimally impacted conditions, is protective of designated uses, and provides management flexibility (USEPA 2000b). However, defining reference condition is not straightforward, so we investigated the effect of several methods for screening reference sites on reference-site-based nutrient criteria. We calculated the 75th percentile of the distributions of nutrient concentrations at reference sites defined by each method in each nutrient ecoregion. Reference-site percentiles are unweighted sample percentiles because both probability and hand-picked sites were screened and combined to define a set of reference sites.

25th percentile of population approach.—Another approach calls for using the 25th percentile of the stream population as a surrogate for reference condition. EPA Office of Water presented draft TP and TN nutrient criteria for each nutrient ecoregion based on the 25th percentile of found data from all seasons in each nutrient ecoregion (e.g., USEPA 2000b). Initial analyses by EPA suggested that the population 25th percentile approximates the reference-site 75th percentile (USEPA 1998). The 25th percentile approach was proposed as a method of last resort if sufficient reference sites were unavailable. We calculated the population 25th percentile with the site-expansion factors from the WSA probability design to estimate true population percentiles. Site-expansion factors are calculated as the inverse of the site-inclusion probability (Herlihy et al. 2000). In these probability surveys, sites were selected from the blue-line stream network on 1:100,000 scale USGS topographic maps. Thus, our estimates refer to wadeable streams depicted on these maps that had water during the summer sampling period.

Modeling natural background concentrations.—Another approach for estimating undisturbed conditions is to model the relationship between nutrient concentrations and human disturbance levels. This model can then be used to predict nutrient concentrations in undisturbed sites. Modeling is advantageous because it eliminates the need to define reference condition or to obtain a lot of data from undisturbed sites. However, the resulting criteria are dependent on the quality and structure of the fitted model. We modeled nutrients with multiple linear regression as a function of disturbance for each nutrient ecoregion in the PNW. We then calculated nutrient concentrations at 0 disturbance (model intercept) and assumed that it represented a background (undisturbed) nutrient concentration that could be used as a potential nutrient criterion. We also calculated a 75th percentile analogue for the modeling approach using the model intercept and standard error (SE) and assuming a normal distribution (75th percentile = intercept + 0.67SE) to facilitate comparison to the 75th percentile approach.

We used $\log_{10}(x)$-transformed TN and $\log_{10}(x + 1)$-transformed TP (TP data had 0 values) as dependent variables. We used only disturbance variables as predictors. Disturbance predictors included watershed (road density, population density, land use/land cover) and physical-habitat (riparian disturbance indices) variables. All disturbance predictors had 0 values in the absence of human disturbance, so the model intercept is the predicted nutrient concentration at 0 human disturbance. We deleted predictor variables that were correlated ($r > |0.90|$), and we $\log_{10}(x)$-transformed variables that were skewed. We used an exhaustive search algorithm to find the best 3- to 6-variable regression models and selected the model with the lowest Bayesian Information Criterion value. We checked residual plots and deleted outliers with Cook's distance >0.5.

Reference-site screening methods

For the WSA data set, we examined only the EMAP reference screening approach. For the PNW data set, we examined several approaches. PNW reference sites were relatively scarce in nutrient ecoregions I and III, so we focused our analysis on nutrient ecoregion II where sampled reference sites were more numerous.

EMAP screening.—We used water-chemistry and physical-habitat data to identify the least-disturbed samples sites in each nutrient ecoregion as has been done in previous EMAP studies (Whittier et al. 2007), with the exception that we did not use nutrients as screening criteria. Nutrient ecoregion-specific screening criteria included measures of water chemistry (SO_4, Cl, turbidity, and pH) and physical habitat (bank canopy density, overall riparian disturbance index, % fine substrate, EMAP riparian disturbance index, and EPA rapid bioassessment protocol habitat score). If a site exceeded the nutrient ecoregion criteria for any 1 variable, then it was not considered a reference site by the EMAP method. In the PNW, we also defined reference condition solely on the basis of the overall EMAP riparian disturbance index using screening values of 0 (no human disturbance observed), ≤0.5, and ≤1.5.

Oregon DEQ screening.—Oregon DEQ has developed a reference-site screening process and has applied it to almost all of the Oregon sites in our database. Their approach is to use GIS and site-specific information to characterize human disturbance. Selected reference sites are then used to describe reference condition for a specific region for the purposes of stream and watershed assessment (Drake 2004). We applied 4 potential reference-site screening tools from the DEQ reference-site identification process: 1) top 20% of candidate reference sites in each Omernik level III ecoregion (top 20%); 2) ideal reference sites (Human Disturbance Index [HDI] class A); 3) ideal, good, and marginal reference sites (HDI classes A, B, and C); and 4) reference sites with a forest fragmentation, urban, road density, reach (FURR) score ≤10. The site HDI is the sum of a reach-level score and a GIS disturbance score. The reach-level score is based on the presence and proximity of 30 different human disturbance activities observed during a sampling visit. The GIS score is based on road density, urban and agriculture land use, and forest fragmentation in the watershed from available GIS layers. HDI scores, best professional judgment, and careful review of site data were used to set HDI classes (Drake 2004). The FURR score is taken only from the HDI GIS evaluation.

Orthophotograph screening.—As part of the EMAP-West analysis, site condition was examined using a Rapid Fine Screening (RFS) process based on GIS and examination of orthophotographs (Stoddard et al. 2005). Sites were scored on an integer scale (0–10), with 0 indicating no sign of human disturbance in the reach. We used 2 different RFS levels (RFS = 0, RFS = 0–3) to identify 2 sets of reference sites.

Biological screening.—During the EMAP survey, ecological condition of each site was assessed based on observed macroinvertebrate assemblages. We used 2 metrics (Hilsenhoff Biotic Index [HBI] and Ephemeroptera, Plecoptera, Trichoptera [EPT] richness) that are commonly used for bioassessment and in indices of biotic integrity (IBI) to identify sites in good condition based on biology. EPT richness is 1 component of the macroinvertebrate IBI developed for the WSA in the PNW (Stoddard et al. 2008). EPT richness is calculated as the sum of the number of mayfly, stonefly, and caddisfly taxa. EPT taxa are generally intolerant to stress, so high EPT scores indicate least-disturbed sites. HBI is the abundance-weighted average of all taxon tolerance values at a site, where taxon tolerances have been defined with respect to an organic pollution gradient (Hilsenhoff 1987). Low numbers indicate intolerant taxa and (presumably) undisturbed conditions. We chose 2 cut-off values for each index in each nutrient ecoregion by best professional judgment to cover a range of possible reference conditions.

Stream typology development

We used stream typology analysis as an independent way to identify groups of least-disturbed streams that have similar expectations for undisturbed nutrient concentrations. We did the analysis only with EMAP-screened reference sites in the PNW. We used regression tree analysis for both TP and TN to identify groups of sites related to each nutrient. We used 197 reference sites that had no missing values for the selected group of predictor variables. We dropped 2 sites that had very high nutrient concentrations (TP > 100 μg/L or TN > 1000 μg/L) that we suspected were not least disturbed. For predictors, we chose natural variables that were available and that we expected to be related to undisturbed nutrient concentrations. These variables included precipitation, runoff, elevation, stream slope, bank canopy density, air temperature, canopy cover, conductivity, acid neutralizing capacity (alkalinity), substrate size, % hardwood forest in watershed, and % fast- and slow-water habitat. We also included Omernik level III ecoregion as a predictor. We determined the size of the tree with 100 repeated 10-fold cross-validations of the data and selected the number of nodes based on the mode of the tree sizes from the minimum deviance trees, i.e., those trees with the smallest cross-validation relative error. We analyzed differences in TP and TN among sites in the tree node classes with 1-way analysis of variance (ANOVA). We used Bonferroni correction to adjust pairwise comparison p values for multiple comparisons. We applied the tree classification to all sites in our data set. We calculated the extent of each class in terms of stream length in the PNW from the probability sites and the 75th percentile of all reference sites in the class (a potential nutrient criterion for the class).

Results

National

Geographic area varied widely among nutrient ecoregions (Fig. 1). Three nutrient ecoregions (II, IX, and XI) accounted for almost 60% of the total wadeable stream length in the continental USA (Table 1). Three other nutrient ecoregions were very small (I, XII, and XIII) and contained <1% of the total wadeable stream length. Because of their small extent, these nutrient ecoregions had too few sample sites for us to make reliable estimates of their potential nutrient criteria, so we excluded them from our national analyses.

Distributions of TP (Fig. 3) and TN (Fig. 4) concentrations varied widely among the nutrient ecoregions. Median TP concentrations were highest in nutrient ecoregions V and X and lowest in nutrient ecoregions II and XI. Patterns in TN population percentiles among nutrient ecoregions were similar to those observed for TP, but the median TN concentrations were highest in nutrient ecoregion VI and lowest in nutrient ecoregion II.

Potential nutrient ecoregion TP and TN criteria values derived from WSA population 25th percentiles

TABLE 1. Potential total P (μg/L) and total N (μg/L) nutrient criteria estimated by Environmental Protection Agency (EPA) found data sample 25th percentile, Wadeable Stream Assessment (WSA) population 25th percentile, WSA Environmental Monitoring and Assessment Program (EMAP)-screened reference-site 75th percentiles in national nutrient ecoregions. See Fig. 1 for nutrient ecoregion names. — indicates sample size too small to make reliable estimates.

	Total P (μg/L)			Total N (μg/L)					
Ecoregion	EPA found 25th	WSA population 25th	WSA reference 75th	EPA found 25th	WSA population 25th	WSA reference 75th	Estimated stream length (km)	No. probability sites	No. reference sites
I	47.0	—	—	310	—	—	2132	4	0
II	10.0	3.0	19.0	120	72.5	148	202,900	526	262
III	21.9	10.4	40.0	380	180	290	41,130	175	89
IV	23.0	18.9	86.8	560	443	926	25,780	107	71
V	67.0	34.4	107.0	880	989	1190	21,650	33	18
VI	76.3	65.8	181.0	2180	1860	2500	101,600	95	53
VII	33.0	17.0	—	540	581	—	85,260	60	7
VIII	10.0	6.8	10.2	380	268	388	126,600	75	40
IX	36.6	20.4	60.1	690	331	681	273,900	157	50
X	128[a]	147.0	—	760	917	—	11,880	20	3
XI	10.0	3.9	17.7	310	160	294	163,700	119	42
XII	40.0	—	—	900	—	—	3419	4	3
XIII	—	—	—	—	—	—	—	0	0
XIV	31.3	22.7	—	710	624	—	24,090	17	7
National total							1,084,000	1392	645

[a] Suspect value in EPA table footnote

and EMAP-screened WSA reference-site 75th percentiles were compared with EPA found data population 25th percentiles (Table 1). TP criteria from all methods were highly correlated ($r > 0.9$). EPA found data population 25th percentiles were higher than WSA population 25th percentiles for all but 1 nutrient ecoregion (Table 1). WSA reference-site 75th percentiles were higher than EPA found data and WSA population 25th percentiles in all nutrient ecoregions, often by factors of 2 to 4. TN criteria from all methods were highly correlated ($r > 0.95$). EPA found data population 25th percentiles were similar to WSA population 25th percentiles in many nutrient ecoregions, but differed by a factor of 2 in nutrient ecoregions III, IX, and XI (Table 1). WSA reference-site 75th percentiles were higher than WSA population 25th percentiles in all nutrient ecoregions.

We examined the relationship between reference-site nutrient concentrations and stream size (based on watershed area) to see whether different nutrient criteria might be needed in different stream size classes. Reference-site nutrient concentrations generally were not correlated with stream size within nutrient ecoregions except in nutrient ecoregions IV (TN, $r = 0.33$, $p < 0.05$) and VI (TP, $r = 0.40$, $p < 0.05$).

PNW

Reference-site screening methods.—Different approaches for defining reference conditions yielded different criterion values in nutrient ecoregion II. TP reference-site 75th percentiles ranged from 11 to 30 μg/L, and TN reference-site 75th percentiles ranged from 88 to 480 μg/L (Table 2). Reference-site 75th percentiles for EMAP-screened PNW sites were 20 μg/L for TP and 165 μg/L for TN. The lowest TP and TN reference-site 75th percentiles were associated with orthophotograph screening. The highest TP and TN reference-site 75th percentiles were associated with Oregon DEQ screening with the higher riparian disturbance index. The various biological screening methods yielded very similar reference-site 75th percentiles (TP: 19–21 μg/L; TN: 133–149 μg/L).

In nutrient ecoregion II, reference-site TP and TN concentrations varied strongly across Omernik level III ecoregions (ANOVA, TP: $F = 19.3$, $p < 0.001$; TN: $F = 12.0$, $p < 0.001$; $n = 200$ EMAP-screened reference sites; Fig. 5). Median TP (2.0 μg/L) and TN (69 μg/L) were much lower in the Omernik level III ecoregion 77 (Northern Cascades) than in the other Omernik level III ecoregions. Median TN (305 μg/L) was much higher in the Omernik level III ecoregion 1 (Coast Range) than in other Omernik level III ecoregions.

Stressor–nutrient modeling.—In nutrient ecoregion I, TP was most strongly related to Cl, annual precipitation, and substrate diameter, whereas TN was most strongly related to % agriculture, % urban, and population density (Table 3). In nutrient ecoregion II, TP was not strongly related to any of the tested

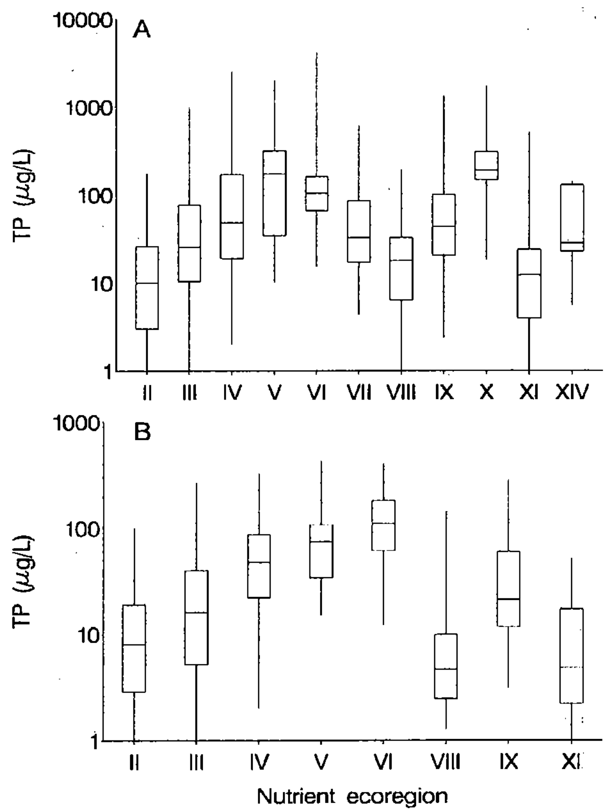

FIG. 3. Box-and-whisker plots for total P (TP) concentrations in all sites based on weighted population estimates (A) and Environmental Monitoring and Assessment Program (EMAP)-screened reference sites (B) in each nutrient ecoregion. Lines in boxes are medians, ends of boxes are quartiles, and whiskers show 1st to 99th percentiles. Only nutrient ecoregions with >10 sites were plotted. See Fig. 1 for nutrient ecoregion names.

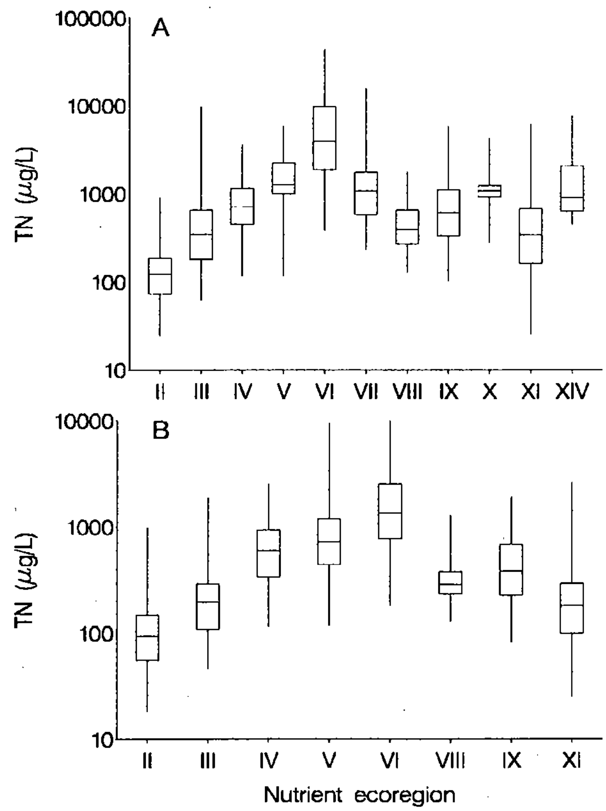

FIG. 4. Box-and-whisker plots for total N (TN) concentrations in all sites based on weighted population estimates (A) and Environmental Monitoring and Assessment Program (EMAP)-screened reference sites (B) in each nutrient ecoregion. Lines in boxes are medians, ends of boxes are quartiles, and whiskers show 1st to 99th percentiles. Only nutrient ecoregions with >10 sites were plotted. See Fig. 1 for nutrient ecoregion names.

environmental variables, whereas TN was most strongly related to Cl, elevation, and February air temperature. In nutrient ecoregion III, TP was strongly related to % agriculture, SO_4, and Cl, % fast water, elevation, February air temperature, and annual precipitation, whereas TN was strongly related to % agriculture, population density, SO_4, Cl, and % fast water.

In nutrient ecoregion I, both the TP and TN multiple regression models with disturbance variables had $r^2 = 0.32$ (Table 4). Watershed % agriculture was a significant predictor variable of both TP and TN. Examination of the model predicted vs observed plots suggested that both models underpredicted nutrients, especially TN, at high concentrations. Back-transformed model intercepts gave the predicted undisturbed TP as 37.5 µg/L and the predicted undisturbed TN as 158 µg/L.

In nutrient region II, the multiple regression models had very low r^2 values (0.18–0.20). Watershed % barren land, road density, and riparian disturbance (roads and nonagriculture) were predictor variables in both TP and TN models (Table 4). However, these models had a narrow prediction range and lacked strong predictive ability. Back-transformed model intercepts gave the predicted undisturbed TP as 10 µg/L and the predicted undisturbed TN as 93 µg/L.

In nutrient ecoregion III, the multiple regression models accounted for more variability ($r^2 = 0.49$–0.60) than in nutrient ecoregions I and II (Table 4). In nutrient ecoregion III, watershed % barren land cover and riparian road disturbance were significant predictor variables for both TP and TN (Table 4). Back-transformed model intercepts gave the predicted undisturbed TP as 28.5 µg/L and the predicted undisturbed TN as 205 µg/L.

TABLE 2. Sample percentiles for total P (TP) and total N (TN) at reference sites defined by different screening methods in nutrient ecoregion II in the Pacific Northwest. See *Reference-site screening methods* for details. EMAP = Environmental Monitoring and Assessment Program; DEQ = Department of Environmental Quality; HDI = Habitat Disturbance Index; FURR = forest fragmentation, urban, road density, reach index; RFS = Rapid Fine Screening; HBI = Hilsenhoff Biotic Index; EPT = Ephemeroptera, Plecoptera, Trichoptera.

		TP (μg/L)		TN (μg/L)	
Screening method	No. of sites	Median	75th percentile	Median	75th percentile
EMAP methods					
EMAP screen	200	10	20	111	165
Riparian disturbance = 0	138	10	20	111	175
Riparian disturbance ≤ 0.5	275	12	29	116	180
Riparian disturbance ≤ 1.5	513	15	30	125	216
Oregon DEQ methods					
DEQ top 20%	80	20	30	170	330
HDI class A	71	20	30	310	480
HDI class A–C	158	20	30	222	390
FURR ≤ 10	32	20	30	162	340
Orthophotographic methods					
RFS = 0	44	2.9	11	67.5	88
RFS = 0–3	96	4.0	17	67.5	108
Biological methods					
HBI < 5	98	7.0	19	66	133
HBI < 4	25	9.7	21	55	133
EPT richness > 10	132	10	20	110	149
EPT richness > 15	48	10	20	114	147

Criterion comparison.—Potential nutrient criteria based on WSA population 25th percentiles and EPA found data population 25th percentiles agreed only weakly (Table 5). For TN, WSA population 25th percentiles were lower than EPA found data population 25th percentiles in all 3 nutrient ecoregions. For TP, WSA population 25th percentiles were lower than EPA found data population 25th percentiles in nutrient ecoregions I and II. Potential nutrient criteria based on WSA reference-site 75th percentiles were higher than potential nutrient criteria based on either 25th percentile approach in almost all comparisons except for TN in nutrient ecoregion III. The regression model estimates of potential nutrient criteria based on the model intercepts and the 75th-percentile analogue of the intercept were always lower than the WSA reference-site 75th percentile.

Stream–nutrient typology.—The regression tree for TP had 5 nodes, or class delineations and explained 46% of the deviance in TP (Fig. 6). The 1st breakpoint was based on long-term annual runoff above/below 0.343 m/y and separated humid from arid areas. The arid group was separated into 2 classes by elevation above/below 1196 m. Streams in the low elevation/arid class (class 5 in Fig. 6) had the highest TP concentrations (Fig. 7). The humid group was separated into 3 classes, first by Omernik level III ecoregion, and then by acid neutralizing capacity (ANC). Humid streams in class 1 were those in Omernik level III ecoregions 15, 16, 17, or 77 and had the lowest TP concentrations. Humid streams in classes 2 and 3 were in other Omernik level III ecoregions and were separated by an ANC value above/below 521 μeq/L. TP differed significantly among the 5 classes (1-way ANOVA, $F = 59.9$, $p < 0.0001$), and all classes except 3 and 4 were significantly different from each other ($p < 0.05$). Arid streams made up 29% of the total stream length in the region, and they were about evenly divided between high- and low-elevation classes (Table 6). Potential nutrient criteria based on WSA reference-site 75th percentiles ranged from 4.0 μg/L in class 1 to 76 μg/L in class 5. Classes 3 and 4 had similar nutrient criteria (32–33 μg/L) (Table 6).

The regression tree for TN explained 48% of the deviance in TN and had 7 nodes (Table 6). The 1st breakpoint divided the sites into 2 groups based on Omernik level III ecoregion. Sites in Omernik level III ecoregions 1, 2, and 3 and nutrient ecoregion III were in 1 group and sites in the other Omernik level III ecoregions were in a 2nd group. The 1st Omernik level III ecoregion group was separated into sites with >9% hardwood forest (class 7), and 2 classes with lower % hardwood forest, which were separated on the basis of substrate size (classes 5 and 6). The high % hardwood forest sites had the highest TN concentrations in the PNW (Table 6). The 2nd Omernik level III ecoregion

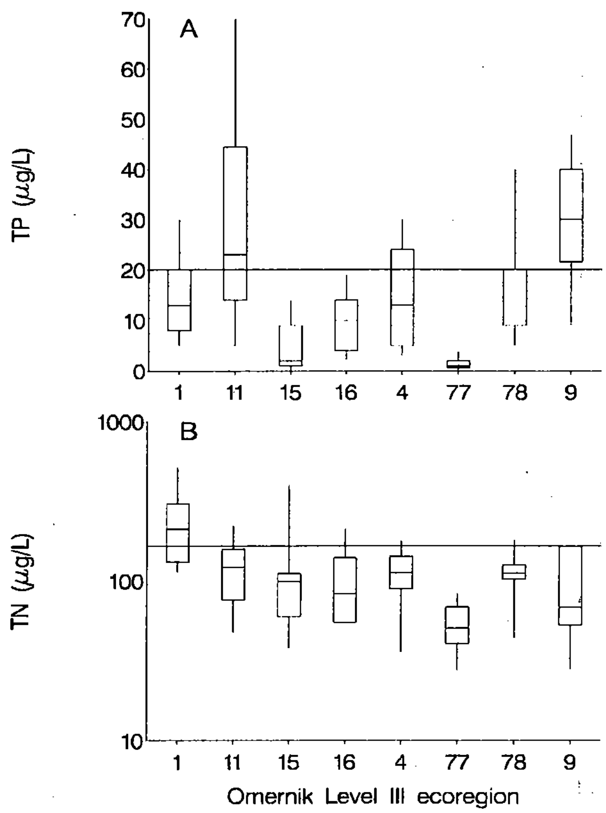

FIG. 5. Box-and-whisker plots for total P (TP) (A) and total N (TN) (B) in Environmental Monitoring and Assessment Program (EMAP)-screened reference sites in Omernik level III ecoregions in nutrient ecoregion II. Omernik level III ecoregions 2 and 17 were not plotted because of low sample size. Lines in boxes are medians, ends of boxes are quartiles, and whiskers show 10th to 90th percentiles. Horizontal reference lines show the nutrient ecoregion II reference-site 75th percentiles. See Fig. 2 for Omernik level III ecoregion names.

group was separated into a class of sites in Omernik level III ecoregions 9 and 77 (class 1) that had the lowest TN in the PNW. The remaining sites were separated into 3 classes on the basis of elevation (above/below 436 m) and then by presence of hardwoods (classes 2–4). The difference in TN among classes was less than the difference in TP among classes, but TN differed significantly among the 7 classes (1-way ANOVA, $F = 32.8$, $p < 0.0001$). TN nutrient criteria were similar for classes 1 and 2, which had criteria significantly lower than criteria for all other classes. TN nutrient criteria did not differ among classes 3, 4, and 5, which had TN criteria that were significantly lower than criteria for classes 6 and 7. TN nutrient criteria did not differ between classes 6 and 7. Class 2 accounted for the largest extent (38%) of stream length in the region (Table 6). Nutrient criteria based on WSA reference-site 75th percentiles ranged from 79 μg/L (class 1) to 509 μg/L (class 7).

Discussion

Setting nutrient criteria

The potential nutrient criteria derived from the different methods were highly correlated in the sense that high criterion values were found consistently in the same nutrient ecoregions and low criterion values were found consistently in the same nutrient ecoregions. However, in terms of setting specific criteria, different approaches yielded very different values (Tables 1, 2, 6, 7). EPA draft criteria derived from found data population 25th percentiles were almost always higher than WSA population 25th percentiles. Found data often focus on problem areas and are not necessarily representative of a region. Thus, found data are likely to have higher nutrient levels than would be observed in the true population. In general, assuming that found data can be used to represent true population statistics is dangerous. Paulsen et al. (1998) gave examples from 4 aquatic case studies in which large discrepancies in regional assessments of condition were found between probability-based survey results and found-data compilations. In addition, the degree to which found data match the population of interest depends greatly on the objectives of the found-data study. Thus, a criterion based on found data will vary depending on the objectives of the data compilations used to define the percentile. Large data compilations from multiple surveys with varying objectives produce percentiles that are virtually uninterpretable.

USEPA (1998) suggests that the population 25th percentile is comparable with the reference-site 75th percentile. In our analyses, the 2 approaches yielded very different potential nutrient criteria in most nutrient ecoregions. In the PNW, WSA population 25th percentiles were not similar to WSA reference-site 75th percentiles. Potential criteria based on WSA reference-site 75th percentiles were higher than either the WSA population 25th percentile or EPA found data population 25th percentile everywhere except for TN in nutrient ecoregion III (Table 5). Nationwide, except for nutrient ecoregion VIII, WSA reference-site 75th percentiles for TP were 1.5 to 4× higher than EPA found data population 25th percentiles and 3 to 6× larger than WSA population 25th percentiles (Table 1). The relationship between potential criteria from WSA reference sites and EPA found data was less biased for TN (Table 1). WSA reference-site 75th percentiles for TN were always higher than WSA population 25th

TABLE 3. Spearman rank correlation coefficients for total P (TP) and total N (TN) vs environmental variables by nutrient ecoregion (I, II, and III) in the Pacific Northwest. See Fig. 1 for nutrient ecoregion names. Correlation coefficients > 0.5 are shown in bold.

	TP			TN		
Variable	I	II	III	I	II	III
Watershed metrics						
% agriculture	0.28	0.20	**0.62**	**0.61**	0.19	**0.65**
% urban	0.29	−0.03	0.23	**0.53**	0.03	0.33
Population density	0.48	0.20	0.47	**0.63**	0.33	**0.55**
Road density	0.12	0.19	0.29	−0.07	0.32	0.09
Watershed area	−0.04	0.04	0.43	−0.01	−0.05	0.44
Water chemistry						
Cl	**0.55**	0.29	**0.53**	0.39	**0.59**	**0.63**
SO_4	0.46	0.03	**0.51**	0.41	0.17	**0.58**
Physical habitat						
Substrate diameter	**−0.53**	−0.40	−0.22	−0.19	−0.19	−0.29
% fast water	−0.37	−0.19	**−0.52**	−0.06	−0.31	**−0.65**
Channel slope	−0.21	−0.19	−0.38	−0.11	−0.27	−0.36
Stream width	−0.41	−0.24	−0.21	−0.26	−0.05	−0.06
Location/climate						
Elevation	−0.29	−0.08	**−0.60**	−0.42	**−0.51**	−0.49
February air temperature	0.06	0.18	**0.57**	0.18	**0.50**	0.39
August air temperature	0.24	0.24	0.38	0.40	0.24	0.40
Annual precipitation	**−0.58**	−0.33	**−0.55**	**−0.50**	0.15	−0.49

percentiles but usually only by a factor of 1.5 to 2×. Suplee et al. (2007) also noted that matches between reference-site and general population percentiles were highly variable. They found that the reference-site 75th percentiles in Montana ecoregions corresponded to general population percentiles ranging from 4th to 97th. Overall, we found no support for the idea that population 25th percentiles can be used as surrogate reference-site 75th percentiles.

The 25th percentile, whether derived from a probability survey or found data, is a flawed way to set nutrient criteria because it is a moving target that can change over time with changing human nutrient use and environmental practices. A probability survey is a definable and consistent way to find a true population 25th percentile, but it treats all nutrient ecoregions as equal, regardless of their level of disturbance. Moreover, by definition, use of the population 25th percentile means that 75% of the sites in any region, regardless of how the region is defined, fail to meet the criterion.

Setting criteria based on natural background levels using undisturbed reference sites seems to be a much

TABLE 4. Multiple regression models used to predict stream total P (μg/L) (TP) and total N (μg/L) (TN) concentrations within the 3 nutrient ecoregions in the Pacific Northwest. See Fig. 1 for nutrient ecoregion names. RD = Riparian Disturbance Index and the word that follows refers to specific disturbance types (see *Data collection* for details). %barren, %developed, %agriculture, road density, and population density are based on watershed landuse data. SE = standard error.

Ecoregion	Regression models	Model r^2	Intercept SE
I (n = 52)	$\log_{10}(TP + 1) = 1.586 + 0.775(RDparks) + 0.667(RDpavement) + 0.579(RDlandfill) - 0.255(RDtotal) + 0.0086(\%agriculture)$	0.32	0.136
	$\log_{10}(TN) = 2.200 + 0.0123(\%agriculture) + 0.0036(\%developed) + 0.661(RDpark)$	0.32	0.148
II (n = 767)	$\log_{10}(TP + 1) = 1.042 + 0.236(RDroad) + 0.243(RDtotal) - 0.274(RDnonagricultural) + 0.0072(road\ density) + 0.0094(population\ density) - 0.0117(\%barren)$	0.20	0.0317
	$\log_{10}(TN) = 1.969 + 0.194(RDroad) + 0.0838(RDtotal) - 0.134(RDnonagricultural) + 0.0272(population\ density) + 0.0099(road\ density) - 0.0057(\%barren)$	0.18	0.0259
III (n = 39)	$\log_{10}(TP + 1) = 1.470 + 0.213(RDtotal) + 1.82(RDbuilding) + 0.404(RDroad) - 1.18(RDlandfill) - 0.0217(\%barren)$	0.49	0.0701
	$\log_{10}(TN) = 2.312 + 0.461(RDroads) - 0.0208(\%barren) + 0.0102(\%agriculture)$	0.60	0.0627

TABLE 5. Potential total P (μg/L) and total N (μg/L) nutrient criteria estimated by Environmental Monitoring and Assessment Program (EMAP) population 25th percentile, EMAP-screened reference-site 75th percentiles (range for all screening methods in parentheses), multiple regressions with riparian disturbance indices as predictor variables (model intercept at 0 disturbance [75th percentile analogue in parentheses]), and Environmental Protection Agency (EPA) found data sample 25th percentile, in the 3 nutrient ecoregions in the Pacific Northwest. See Fig. 1 for nutrient ecoregion names.

Nutrient ecoregion	Population 25th	Reference 75th	Regression model	EPA found 25th
TP				
I	21	70 (21–81)	37.5 (46.5)	47
II	4.0	20 (11–30)	10.0 (10.6)	10
III	33	49 (43–101)	28.5 (31.9)	21.9
TN				
I	189	471 (289–593)	158 (199)	310
II	55.0	165 (88–480)	93.1 (126)	120
III	221	285 (285–801)	205 (226)	380

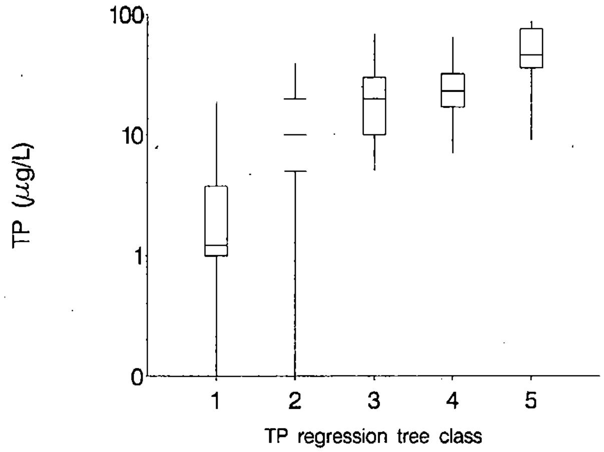

FIG. 7. Box-and-whisker plot for total P (TP) for sites in each TP regression-tree class in the Pacific Northwest (Fig. 6). Lines in boxes are medians, ends of boxes are quartiles, and whiskers show ranges.

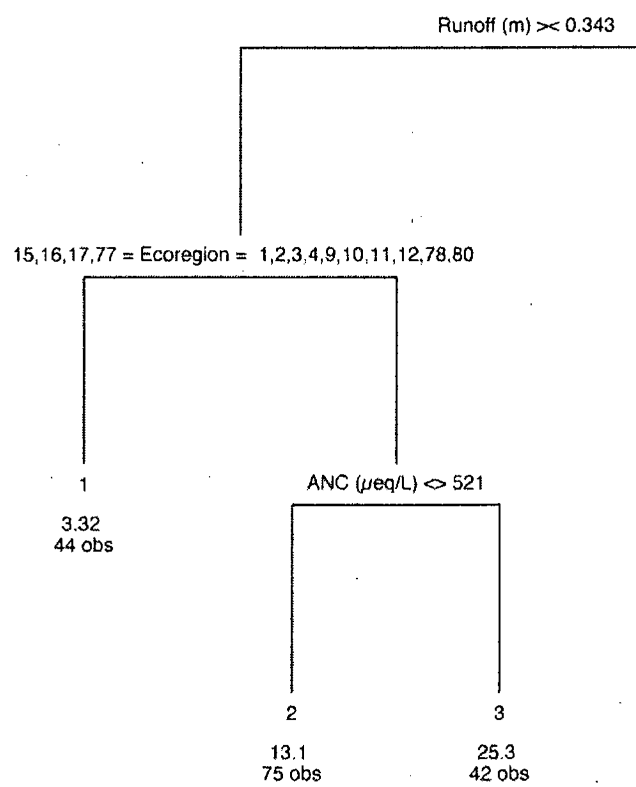

FIG. 6. Regression tree for total P (TP) in Environmental Monitoring and Assessment Program (EMAP)-screened reference sites in the Pacific Northwest. The node or class number is shown immediately below the branch. The value below that is the mean TP concentration (μg/L) for the sites in that class. The tree can be read like a dichotomous key with the equation at each split indicating that the cases with the lower values go left (<>) or right (><) at the split. ANC = acid neutralizing capacity, runoff = 30-y long-term mean annual runoff, obs = observations. Ecoregions are given as Omernik (1987) level III ecoregions numeric codes. See Fig. 2 for ecoregion names.

TABLE 6. Total P (μg/L) (TP) and total N (μg/L) (TN) regression-tree class descriptors, Environmental Monitoring and Assessment Program (EMAP) estimates of wadeable stream length in each class, and potential TP and TN criteria for the class based on EMAP-screened reference-site 75[th] percentiles in the 3 nutrient ecoregions in the Pacific Northwest. Humid = long-term mean annual runoff > 0.343 m, Arid = long-term mean annual runoff < 0.343 m, ANC = acid neutralizing capacity. See Fig. 2 for Omernik level III ecoregion names.

Site class	Description	Stream length (km)	No. reference sites	Reference-site 75[th] percentile
TP				
1	Humid; in Omernik level III ecoregions 15, 16, 17, or 77	27,100 (25%)	46	4.0
2	Humid; not in class 1 Omernik level III ecoregions; ANC < 521 μeq/L	30,000 (27%)	87	20
3	Humid; not in class 1 Omernik level III ecoregions; ANC > 521 μeq/L	20,900 (19%)	51	32
4	Arid; elevation >1196 m	18,100 (16%)	27	33
5	Arid; elevation < 1196 m	13,800 (13%)	13	76
TN				
1	In Omernik level III ecoregions 9, 77	14,300 (13%)	38	79.0
2	In Omernik level III ecoregions 4, 11, 15, 16, 17, or 78; elevation > 436 m and % hardwood forest < 0.07%	41,600 (38%)	40	100
3	In class 2 Omernik level III ecoregions; elevation > 436 m; % hardwood forest > 0.07%	21,200 (19%)	61	145
4	In class 2 Omernik level III ecoregions; elevation < 436 m	4600 (4%)	16	168
5	In Omernik level III ecoregions 1, 2, or 3 or nutrient ecoregion III; % hardwood forest < 9%; substrate diameter > 22 mm (coarse gravel)	10,400 (10%)	32	216
6	In class 5 ecoregions; % hardwood forest < 9%; substrate diameter < 22 mm	7070 (6%)	16	345
7	In class 5 ecoregions; % hardwood forest > 9%	10,200 (9%)	21	509

more defensible approach than using percentiles of found data. A large enough reference-site sample would help control for natural variability and provide realistic field-measurement-based criterion values. The drawbacks to a reference-site approach are that reference site definitions can change over time and good reference sites do not exist in many regions. In a strict sense, pristine sites do not exist at all because atmospheric deposition contains elevated levels of nutrients. A reference-based approach also is heavily dependent on the definition of what constitutes a reference site. Reference-site definitions typically default to a least-disturbed definition (Stoddard et al. 2006). Unfortunately, this definition can be highly variable from region to region and even person to person. Thus, a major problem with the reference approach is use of different standards or benchmarks to compare sites in different regions or classes. Ecore-

TABLE 7. Comparison of total P (TP) and total N (TN) background reference criteria and % population stream length exceeding reference criteria in our study and other national analyses. Background criteria were derived from the Wadeable Stream Assessment (WSA) reference-site 75[th] percentiles, or modeled by Dodds and Oakes (2004; table 7) or Smith et al. (2003; table S3, TN values are from the "with deposition" output and include current atmospheric N contributions).

	TP (μg/L)						TN (μg/L)					
	Background criteria			% stream length exceeding			Background criteria			% stream length exceeding		
Ecoregion	WSA	Dodds and Oakes (2004)	Smith et al. (2003)	WSA	Dodds and Oakes (2004)	Smith et al. (2003)	WSA	Dodds and Oakes (2004)	Smith et al. (2003)	WSA	Dodds and Oakes (2004)	Smith et al. (2003)
II	19	45	20	32	12	32	148	479	210	36	5	21
III	40	151	30	34	11	45	290	918	110	57	14	90
IV	87	59	70	36	45	39	926	659	210	34	56	87
V	107	23	70	62	78	64	1190	566	510	56	87	91
VI	181	23	60	24	97	77	2500	215	620	63	100	93
VIII	10	28	20	67	26	45	388	589	280	51	36	71
IX	60	31	50	39	60	46	681	370	280	44	69	82
XI	18	43	20	40	13	36	294	1102	290	53	11	54
All				39	40	45				47	42	65

gional variation in reference-site quality was a major issue in the analysis of WSA macroinvertebrate data (Herlihy et al. 2008).

Defining reference sites

We defined reference sites for nutrient ecoregion II in the PNW with a variety of methods (Table 2). Biological screening methods based on 2 macroinvertebrate metrics with 2 different cut-off values gave very consistent reference-site 75th percentiles for both TP (19–21 μg/L) and TN (133–149 μg/L). Wang et al. (2007) noted a strong relationship between nutrient levels and many fish and macroinvertebrate metrics in Wisconsin streams and suggested that nutrient–biota relationships be used to validate and refine nutrient criteria. In our data set, nutrient–HBI and nutrient–EPT richness plots showed no break points or thresholds. The correlations were statistically significant, but highly scattered, so we did not attempt to use these relationships to define specific nutrient criteria.

The lowest nutrient concentrations were seen in nutrient ecoregion II reference sites defined by the most restrictive (least human influence) orthophotograph screening. As reference criteria were relaxed in the orthophotograph and EMAP riparian disturbance screening, reference-site nutrient concentration percentiles increased (Table 2). These results strongly suggest that a human-disturbance gradient exists in least-disturbed sites. Thus, reference-site definitions do significantly affect potential nutrient criterion values. We used our more restrictive definitions of reference to calculate reference-site 75th percentile potential nutrient criteria of 11 to 20 μg/L for TP and 88 to 165 μg/L for TN in nutrient ecoregion II.

Oregon DEQ screening yielded reference sites with higher nutrient percentiles than did other screens. However, Oregon DEQ screening was applied only to Oregon sites, and only data from Oregon reference sites were used to estimate the reference-site 75th percentile. In contrast, the other screening methods were applied to all sites in the PNW, and reference-site 75th percentiles were based on the whole PNW. A strong Omernik level III ecoregion effect was observed across nutrient ecoregion II (Fig. 5), so the restriction to Oregon data might explain some of the variation in reference-site nutrient concentrations among screening methods. Nutrient concentrations were higher in ecoregions in Oregon and lower in ecoregions in Washington and Idaho (Fig. 5). The Omernik level III ecoregion effect also might explain some of the variation in reference-site nutrient concentrations among screening methods (Table 2). The distribution of reference sites across ecoregions varied by screening method because some ecoregions are more disturbed than others.

Setting nutrient criteria from models

Other researchers have investigated setting potential nutrient criteria from reference-site distributions or by modeling natural background concentrations. Suplee et al. (2007) used the 75th percentile of reference sites to develop nutrient criteria for Montana streams. They reported reference-site 75th percentiles (TP: 3–20 μg/L; TN: 90–175 μg/L) for Omernik level III ecoregions in nutrient ecoregion II that are within the range of values we found with WSA and PNW data. Suplee et al. (2007) reported higher reference-site 75th percentiles for TP in Montana streams in nutrient ecoregions IV (170 μg/L) and V (140 μg/L) than we observed with WSA data in nutrient ecoregions IV (87 μg/L) and V (107 μg/L). However, for TN, Montana reference-site 75th percentiles (1120 μg/L) were virtually identical to WSA reference-site 75th percentiles (1190 μg/L) in nutrient ecoregion V and somewhat higher (1300 μg/L) than WSA reference-site 75th percentiles (926 μg/L) in ecoregion IV. The 75th percentiles for 85 sites in undeveloped watersheds across the USA (Clark et al. 2000) were 37 μg/L for TP and 500 μg/L for TN, values that fall within the range of WSA values from national nutrient ecoregions.

Two studies have calculated potential nutrient criteria for all of the national nutrient ecoregions based on modeling. Dodds and Oakes (2004) constructed multiple regression models of TN and TP vs various anthropogenic landuse categories and calculated reference concentrations as the intercept of the model when anthropogenic activity was 0. Smith et al. (2003) developed an empirical model (SPARROW) to calculate background yield of watershed TP and TN based on runoff, basin size, current atmospheric N deposition, and region-specific factors. Model output was used to estimate the 75th percentile of predicted natural background concentrations for each nutrient ecoregion. We compared WSA reference-site 75th percentiles for TP and TN with values modeled by Dodds and Oakes (2004) and Smith et al. (2003) (Table 7). We used the WSA population estimates to calculate the % of stream length in each nutrient ecoregion that exceeded criteria from each of the 3 approaches (Table 7). In the 8 nutrient ecoregions with enough WSA reference sites to do this analysis, 39% of total stream length exceeded WSA TP criteria and 47% exceeded TN criteria. These 8 ecoregions account for 88% of the stream length in the conterminous states.

TP criteria based on the Smith et al. (2003) model were similar to those from the WSA, except in nutrient ecoregions V and VI (Table 7). TP criteria based on the Dodds and Oakes (2004) model were either much higher or much lower, depending on the nutrient ecoregion, than those from the WSA. The WSA TP criteria for nutrient ecoregions V and VI were substantially higher than TP criteria from the Smith et al. (2003) or Dodds and Oakes (2004) models. WSA reference sites in nutrient ecoregions V and VI might be least disturbed for those ecoregions, but they probably are influenced by anthropogenic nutrient additions. The TP criterion based on the Dodds and Oakes (2004) model for nutrient ecoregion VI is very low (23 μg/L) and virtually all (97%) of the stream length in that region exceeds that criterion. TN criteria based on the Smith et al. (2003) model are similar to those from the WSA in mountainous nutrient ecoregions II and XI. WSA TN criteria are much higher than those based on the Smith et al. (2003) model in the 6 lower-elevation nutrient ecoregions. In these 6 lower elevation ecoregions, 71 to 93% of total stream length exceeds the Smith et al. (2003) TN criteria (Table 7).

Our disturbance-based regression model estimates of potential nutrient criteria in the PNW were lower than the reference-site 75th percentile regardless of whether we used model intercepts or the 75th percentiles of model intercepts. The r^2 value was much lower for the nutrient ecoregion II model than for nutrient ecoregion I or III models, probably because the disturbance signal in nutrient ecoregion II was weak. In the case where the disturbance regression model $r^2 = 0$, the regression-model criterion (intercept) becomes the population mean. Suplee et al. (2007) compared their reference-site 75th percentiles to nutrient criteria obtained from 5 different modeling approaches. Modeled nutrient criteria matched, on average, the 86th percentile of the reference-site distribution (Suplee et al. 2007), a result that lends support to the use of the reference-site 75th percentile as a basis for setting criteria. Robertson et al. (2006) modeled reference-site nutrient concentrations with a multiple regression model based on watershed % agriculture and urban land cover for streams in the Upper Midwest. They found reference (0 disturbance) TP concentrations of 84 μg/L in nutrient ecoregion VI and 12 to 19 μg/L in nutrient ecoregions VII, VIII, IX, and XI. These values agree with WSA reference levels in ecoregions VIII and IX, but are lower than WSA values in ecoregions VI and IX.

Variability in nutrient concentrations among ecoregions

In the PNW, we observed wide variation in reference-site nutrient concentrations among Omernik level III ecoregions within nutrient ecoregion II (Fig. 5). Reference-site 75th percentiles for TP were 2 μg/L in Omernik level III ecoregion 77 vs 45 μg/L in the Omernik level III ecoregion 11. These results suggest that a single nutrient criterion for all wadeable streams in nutrient ecoregion II is inappropriate. Natural factors cause high variability in reference nutrient concentrations among streams within the national nutrient ecoregions (Smith et al. 2003, Dodds and Oakes 2004, Robertston et al. 2006). Modeled background nutrient concentrations in streams in some nutrient ecoregions ranged over an order of magnitude, in large part because of differences in local runoff (Smith et al. 2003). Wickham et al. (2005) observed that landuse classes explained 3 to 6× more variation in nutrient chemistry than did nutrient ecoregions and suggested that land-cover composition also be used to guide development of nutrient criteria. Overall, these results indicate that the 14 national nutrient ecoregions are too coarse for setting nutrient criteria.

Previous work in PNW lakes and reservoirs also suggested that national nutrient ecoregions and even Omernik level III ecoregions were too coarse for setting nutrient criteria. A typology of Omernik level III ecoregion 1 lakes includes coastal dune lakes, dystrophic bog lakes, and mountain lakes that all have different expectations for nutrient levels and processing (Vaga et al. 2005). A 3-class typology for reservoirs in the PNW is based on ionic strength and turbidity and is strongly related to nutrient levels (Vaga et al. 2006). We used regression-tree analysis to develop a typology that was responsive to natural gradients of nutrient concentrations for PNW streams. Runoff, elevation, ANC, forest composition, substrate size, and Omernik level III ecoregion were related to TP or TN concentrations in least-disturbed sites. Robertson et al. (2006) used a type of regression-tree analysis to develop an alternate nutrient criteria classification for the Upper Midwest. They analyzed raw data and residual data with the effects of land use statistically removed and were able to separate small watersheds in the Upper Midwest into 5 relatively homogeneous environmental water-quality zones. Robertson et al. (2006) used modeling and frequency-distribution approaches to develop nutrient criteria for each zone. They suggested, as do we, that this type of modeling yields more appropriate nutrient criteria than does application of single criteria to all sites within large regions.

Nutrient ecoregions were developed to minimize natural variability in nutrient concentrations within groups of sites so that nutrient criteria could be set that applied across all sites within the group. Our analysis and the literature strongly suggest that the 14 national nutrient ecoregions are too coarse to achieve this objective. Advances in setting national nutrient criteria will require a finer-scale stream typology or classification of sites that better controls for natural variations in nutrient concentration.

Acknowledgements

Our work was funded by a grant from US EPA Region 10 (X7-970947-01) and cooperative agreement CR831682-01 between Oregon State University and the US EPA National Health and Ecological Effects Research Laboratory–Western Ecology Division. We thank all the people involved with the EMAP-West and the WSA for insight and for sharing data. We also thank Shannon Hubler, Doug Drake, and Mike Mulvey (Oregon DEQ) for sharing survey data and reference-site classification work.

Literature Cited

CLARK, G. M., D. K. MUELLER, AND M. A. MAST. 2000. Nutrient concentrations and yields in undeveloped stream basins of the United States. Journal of the American Water Resources Association 36:849–860.

DODDS, W. K. 2007. Trophic state, eutrophication and nutrient criteria in streams. Trends in Ecology and Evolution 22: 670–676.

DODDS, W. K., AND R. M. OAKES. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. Limnology and Oceanography: Methods 2:333–341.

DRAKE, D. 2004. Selecting reference condition sites, an approach for biological criteria and watershed assessment. Technical Report WAS04–002. Oregon Department of Environmental Quality, Portland, Oregon. (Available from: http://www.deq.state.or.us/lab/techrpts/docs/WSA04002.pdf)

HERLIHY, A. T., P. R. KAUFMANN, L. REYNOLDS, J. LI, AND G. ROBISON. 1997. Developing indicators of ecological condition in the Willamette Basin – an overview of the Oregon Prepilot Study for EPA's EMAP Program. Pages 275–282 *in* A. Laenen and D. Dunnette (editors). River quality: dynamics and restoration. Lewis Publishers, CRC Press, Boca Raton, Florida.

HERLIHY, A. T., D. P. LARSEN, S. G. PAULSEN, N. S. URQUHART, AND B. J. ROSENBAUM. 2000. Designing a spatially balanced, randomized site selection process for regional stream surveys: the EMAP mid-Atlantic pilot study. Environmental Monitoring and Assessment 63:95–113.

HERLIHY, A. T., S. G. PAULSEN, J. VAN SICKLE, J. L. STODDARD, C. P. HAWKINS, AND L. L. YUAN. 2008. Striving for consistency in a national assessment: the challenges of applying a reference condition approach at a continental scale. Journal of the North American Benthological Society 27:860–877.

HILSENHOF, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20:31–39.

KAUFMANN, P. R., P. LEVINE, E. G. ROBISON, C. SEELIGER, AND D. V. PECK. 1999. Quantifying physical habitat in wadeable streams. EPA 620/R-99/003. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

MOSER, T. J., P. J. WIGINGTON, M. J. SCHUFT, P. R. KAUFMANN, A. T. HERLIHY, J. VAN SICKLE, AND L. S. MCALLISTER. 1997. The effect of riparian areas on the ecological condition of small, perennial streams in agricultural landscapes of the Willamette Valley. EPA/600/R-97/074. Office of Research and Development, US Environmental Protection Agency, Corvallis, Oregon.

OLSEN, A. R., AND D. V. PECK. 2008. Survey design and extent estimates for the Wadeable Stream Assessment. Journal of the North American Benthological Society 27:822–836.

OMERNIK, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118–125.

PAULSEN, S. G., R. M. HUGHES, AND D. P. LARSEN. 1998. Critical elements in describing and understanding our nation's aquatic resources. Journal of the American Water Works Association 34:995–1005.

PECK, D. V., A. T. HERLIHY, B. H. HILL, R. M. HUGHES, P. R. KAUFMANN, D. J. KLEMM, J. M. LAZORCHAK, F. H. MCCORMICK, S. A. PETERSON, P. L. RINGOLD, T. MAGEE, AND M. R. CAPPAERT. 2006. Environmental Monitoring and Assessment Program–Surface Waters Western Pilot Study: field operations manual for wadeable streams. EPA 620/R-06/003. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

ROBERTSON, D. M., D. A. SAAD, AND D. M. HEISEY. 2006. A regional classification scheme for estimating reference water quality in streams using land-use-adjusted spatial regression-tree analysis. Environmental Management 37: 209–229.

ROHM, C. M., J. M. OMERNIK, A. J. WOODS, AND J. L. STODDARD. 2002. Regional characteristics of nutrient concentrations in streams and their application to nutrient criteria development. Journal of the American Water Resources Association 38:213–239.

SMITH, R. A., R. B. ALEXANDER, AND G. E. SCHWARZ. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. Environmental Science Technology 37:3039–3047.

STEVENS, D. L., AND A. R. OLSEN. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278.

STODDARD, J. L., A. T. HERLIHY, D. V. PECK, R. M. HUGHES, T. R. WHITTIER, AND E. TARQUINIO. 2008. A process for creating multimetric indices for large-scale aquatic surveys. Journal of the North American Benthological Society 27:878–891.

STODDARD, J. L., D. P. LARSEN, C. P. HAWKINS, R. K. JOHNSON, AND R. H. NORRIS. 2006. Setting expectations for the ecological condition of running waters: the concept of reference condition. Ecological Applications 16:1267–1276.

STODDARD, J. L., D. V. PECK, A. R. OLSEN, D. P. LARSEN, J. VAN SICKLE, C. P. HAWKINS, R. M. HUGHES, T. R. WHITTIER, G. LOMNICKY, A. T. HERLIHY, P. R. KAUFMANN, S. A. PETERSON,

P. L. RINGOLD, S. G. PAULSEN, AND R. BLAIR. 2005. Environmental Monitoring and Assessment Program (EMAP): western streams and rivers statistical summary. EPA 620/R-05/006. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

SUPLEE, M. W., A. VARGHESE, AND J. CLELAND. 2007. Developing nutrient criteria for streams: an evaluation of the frequency distribution method. Journal of the American Water Resources Association 43:453–471.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1987. Handbook of methods for acid deposition studies: laboratory analysis for surface water chemistry. EPA 600/4-87/026. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1996. National Nutrient Assessment Workshop Proceedings. EPA-822-96-004. Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1998. National strategy for the development of regional nutrient criteria. EPA-8221-R-98-002. Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000a. Nutrient criteria technical guidance manual: rivers and streams. EPA-822-B00-002. Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000b. Ambient water quality criteria recommendations. Information supporting the development of state and tribal nutrient criteria for rivers and streams in nutrient ecoregion III. EPA 822-B-00-016. Office of Water, US Environmental Protection Agency, Washington, DC.

VAGA, R. M., A. T. HERLIHY, R. MILLER, AND M. M. SYTSMA. 2006. A classification of Pacific Northwest reservoirs with respect to nutrient processing. EPA 910-R-06-003. Region 10, US Environmental Protection Agency, Seattle, Washington.

VAGA, R. M., R. R. PETERSEN, M. M. SYTSMA, M. ROSENKRANTZ, AND A. T. HERLIHY. 2005. A classification of lakes in the Coast Range ecoregion with respect to nutrient processing. EPA 910-R-05-002. Region 10, US Environmental Protection Agency, Seattle, Washington.

VAN SICKLE, J., J. BAKER, A. HERLIHY, P. BAYLEY, S. GREGORY, P. HAGGERTY, L. ASHKENAS, AND J. LI. 2004. Projecting the biological condition of streams, under alternative conditions of human land use. Ecological Applications 14:368–380.

WANG, L., D. M. ROBERTSON, AND P. J. GARRISON. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: implications to nutrient criteria development. Environmental Management 39:194–212.

WHITTIER, T. R., J. L. STODDARD, D. P. LARSEN, AND A. T. HERLIHY. 2007. Selecting reference sites for stream biological assessments: best professional judgment or objective criteria. Journal of the North American Benthological Society 26:349–360.

WICKHAM, J. D., K. H. RIITERS, T. G. WADE, AND K. B. JONES. 2005. Evaluating the relative roles of ecological regions and land-cover composition for guiding establishment of nutrient criteria. Landscape Ecology 20:791–798.

Received: 7 March 2008
Accepted: 21 July 2008